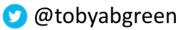


# Scholarship Beyond Books and Journals

How digital is changing how 'human thought' is published

#### **Toby Green**

Co-Founder, Coherent Digital



https://orcid.org/0000-0002-9601-9130

toby.green@coherentdigital.net



1 Hat tip: Filippos Tsimpoglou, Director General National Library of Greece





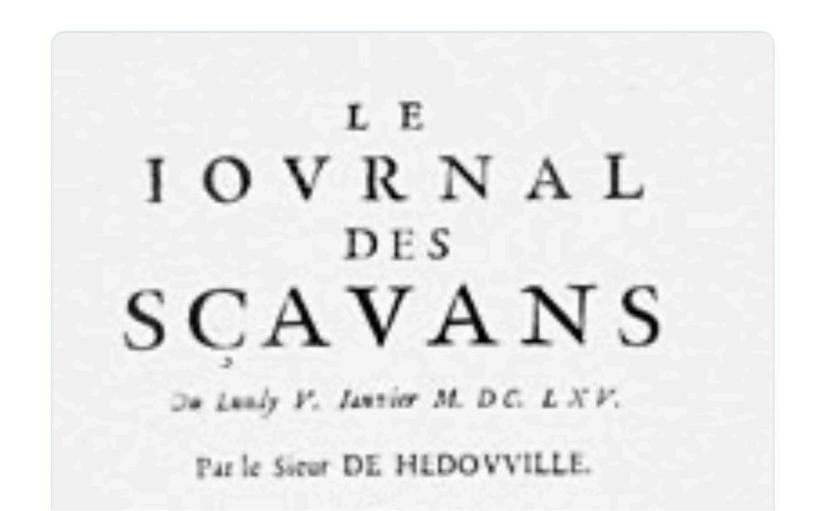
The scholarly record has always comprised \*more\* than books & journals. Curiously, digital makes it harder for the \*more\* to join today's record & much is now missing. Librarians and scholars are worried. A long (you might want to make ) if irst).

#Fiesole2022



1665. The first scholarly journal - a twelve page quarto pamphlet - is published. An industry is born .

https://en.wikipedia.org/wiki/Journal\_des\_s%C3%A7avans





1665-1996. The publisher-library 'complex' evolves a sophisticated system to publish, curate, collect & disseminate 'human thought'. Serving both readers and authors, it has libraries at its centre.





to the end of the print era. A lot of scholarly has been published and carefully archived. (Lest we forget, standing on the shoulders of giants is only possible thanks to the work of librarians.





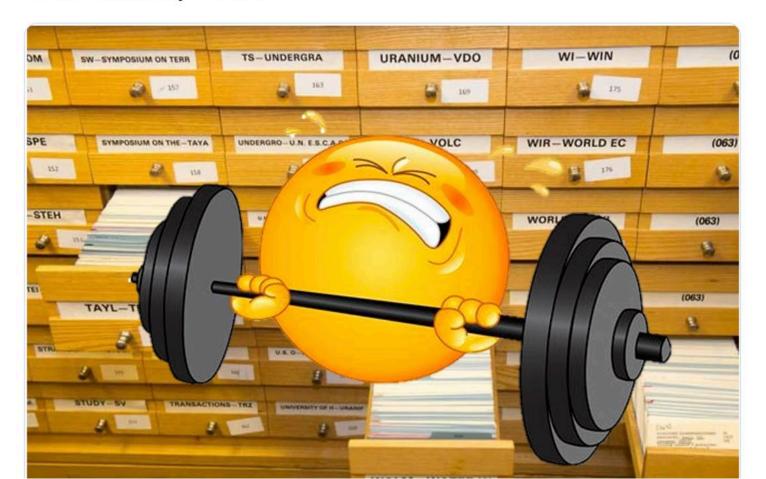
Librarians know that giants need a varied diet - so alongside formally published books and journals librarians also collect #primarysources and #greylit, if they can find it. ("if" doing a lot of work here).





Toby Green wants your feedback on @Policy\_Commons @Toby... Why "if"?

Lacking a publisher (yes, publishers add value), #greylit and #primarysources are challenging for librarians to find, catalogue and index - but thanks to their efforts, this content took its place in the scholarly record.





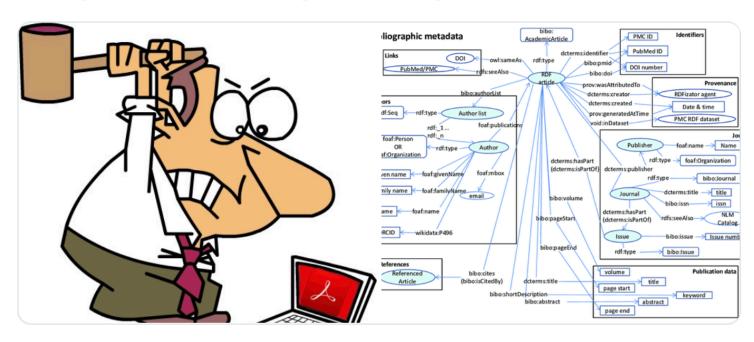
Toby Green wants your feedback on @Policy\_Commons @Toby...
So, the scholarly record's in good shape to enter the digital age?

So, the scholarly record's in good shape to enter the digital age? Right?





The trouble with digital isn't the much-maligned PDF file . . . it's the digital scaffolding - the metadata - needed to create, manage and find the PDF file. Digital scaffolding is way harder to do than analogue. (I'm a ﷺ, so forgive my using unnecessary 'u' & 'e's)





1996-2004. Digital era 1.0. Publishers and librarians move fast to digitise the scholcomm loop, building digital systems (with colourful logos!) at the heart of which is standardised metadata that leads to the PDFs (and other digital stuff too).





But #greylit is left behind. Sure, their producers learn how to produce PDFs, but they don't know how to create metadata (and they don't much care either). So this content becomes harder to find & catalogue. It's excluded from today's scholarly record ...



### **Primary Sources**

**Diaries** 

Photographs Letters Newspapers

Spaachac



#### **Grey Literature**

Pamphlets White Papers

**Reports** 

**Financial statements** 



Toby Green wants your feedback on @Policy\_Commons @Toby...

Digital has effectively put a bouncer on the door - No metadata? No entry!





# **Toby Green wants your feedback on @Policy\_Commons** @Toby... Along comes 2004 . . . and a surprise . . .

No, not THIS surprise (but 🤏 Greece, for 🎳 🟆 Euro 2004 🗳 🍆)

... no, I was thinking about ...





Blogs Participation Statement Usability Widgets Recommendation Social Softwareow Videocasting Podcasting Collaboration Perpetual Sets Simplicity Convergence Web 2.0 CSS PayPer Click Mobility Atom NOTINE SVG Ruby on Ralla VC Trust Affiliation OpenAPIs RSS Semantic Web Standardsco Economy OpenID Remixability REST StandardizationThe Long Tail DataDriven Accessibility Microformats Syndication

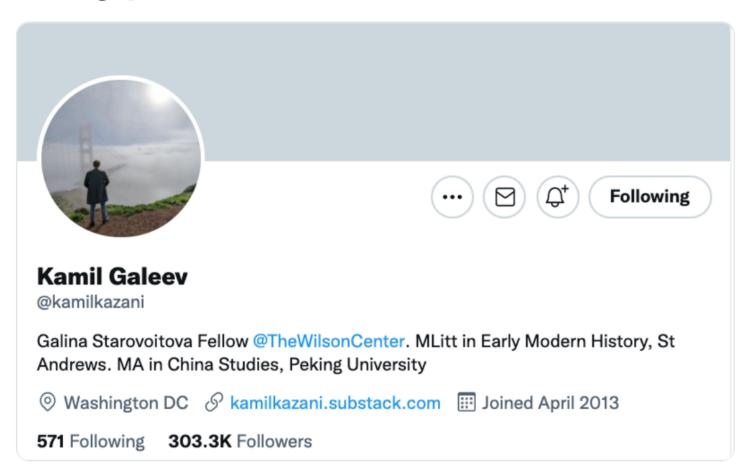


Web 2.0 unleashed a wave of 'user-generated' content - chaotic, creative, and contumacious (aka stubbornly or wilfully disobedient to authority). But what's that got to do with scholcomm and the scholarly record? Surely academics won't publish on social media?



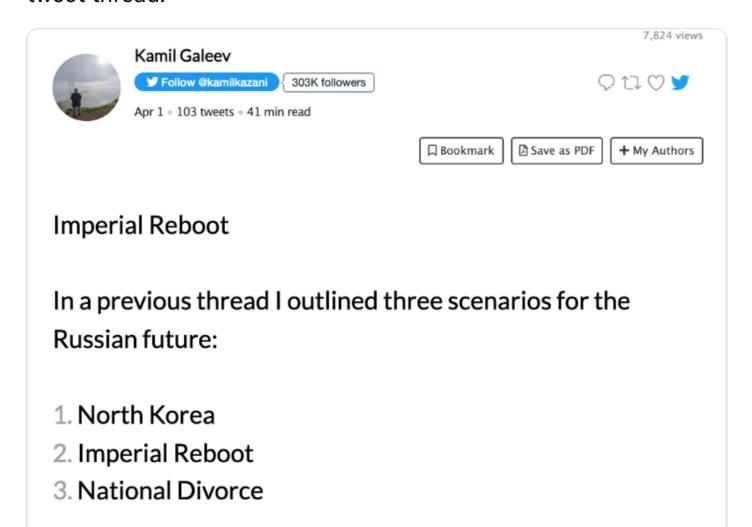


Allow me to introduce Kamil Galeev @kamilkazani - a Galinova Starovoitova Fellow at the Wilson Center and of St Andrews and Peking Universities. He has over 300k followers. He writes cracking s about Ukraine-Russia.





Take this one. It's a 103-tweet and takes ~40 minutes to read. It's packed with links, references and citations - just like a working paper or preprint. Except he chose to publish his 'paper' as a tweet thread.





Or how about Professor Chris Grey @chrisgreybrexit who has been writing a weekly blog about #Brexit since 2016. Currently at Royal Holloway, Professor Grey used to be at Warwick and Cambridge Universities and earned his PhD at Manchester.





Chris Grey's Brexit & Beyond blog runs to over 350 weekly posts (some are very long) and they've racked up nearly 7 million views. Prediction: it will be one of the most important primary sources for studying the history of Brexit. And it's currently on the Blogspot platform  $\widehat{\mathbf{w}}$ .

### **Brexit & Beyond**

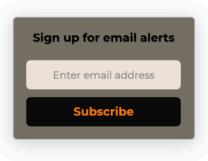
"Best guy to follow on Brexit for intelligent analysis" Annette Dittert, ARD German TV. "By far one of the best analysts of Brexit" Sarah Carey, The Times. "Consistently outstanding analysis of Brexit" Jonathan Dimbleby. "The best writer on Brexit" Chris Lockwood, Europe Editor, The Economist. "A must-read for anyone following Brexit" David Allen Green, FT. "The doyen of Brexit commentators" Chris Johns, Irish Times.

Sunday, 4 September 2016

#### Where we are now

This new blog will comment on the aftermath of the UK Referendum decision to leave the European Union. It starts from the position that this decision was a national catastrophe to which I was completely opposed, and that it was made on the basis of persistent, shameless lies from the Leave campaign which seriously mislead the British people. However, the blog will not attempt to re-argue the case for remaining in the EU or for the reversal of the decision, as I do not think that this is in prospect. Instead, I will analyse the unfolding consequences. As an opening post, I will summarise the current situation and in future posts discuss issues and developments as they emerge.

As the summer ends, serious consideration of the consequences of the vote to leave the EU begins. Thus far, the British Prime Minister Theresa May has been able to get away with the formula 'Brexit means Brexit' and thus avoid the obvious questions



#### **Pages**

- About me and this blog
- Media writings, quotes, appearances, public policy etc
- Comments policy



Or, how about the Intergovernmental Panel on Climate Change @ipcc\_ch which has been in the news this week because they released the third report in their sixth assessment cycle. Its 207 authors are led by Chairs and teams based at Imperial College and Ahmedabad University.



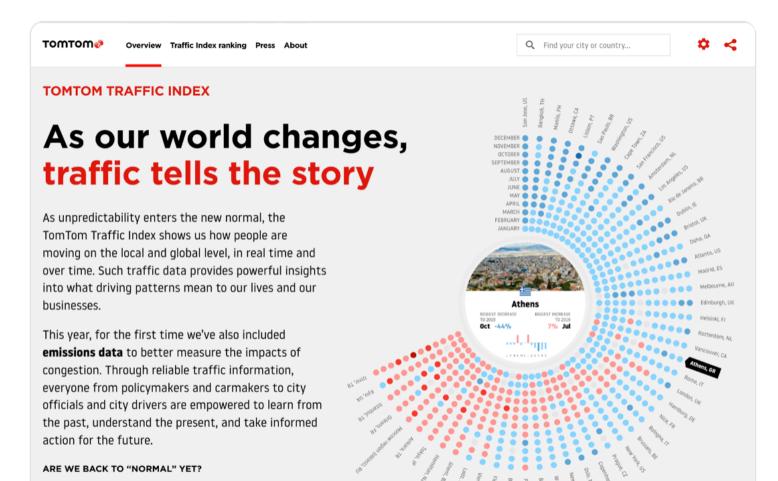


Until 2014, IPCC published their reports as books, formally, with a publisher. Now they simply release them on their website as #greylit. What was formal is now grey. What was easy to collect is now hard. What should be in the scholarly record, is now missing.



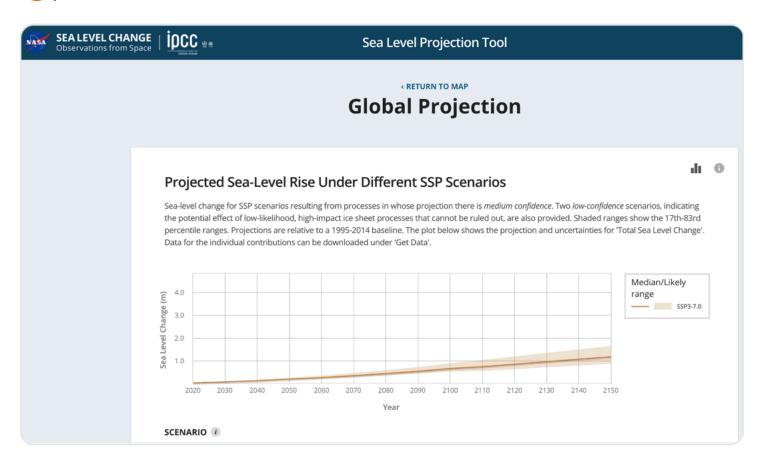


... and then there's stuff that doesn't involve the prose (or academics). Guess what? We all help to build this one! How many libraries have tomtom traffic index catalogued? It's important. Anyone studying #climate, #urbanplanning or post-Covid #socialchange needs it.





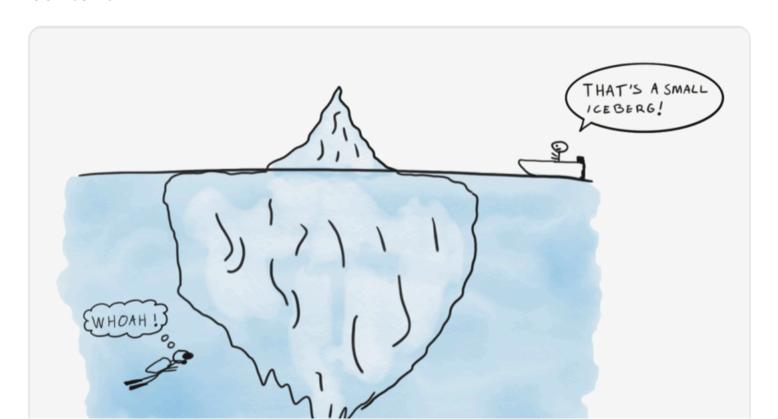
And what about this? IPCC's projections for sea-level rise. You might expect to find this interactive dataset and maps on their website. Except it's not there. It's on NASA's website (obviously (2)).





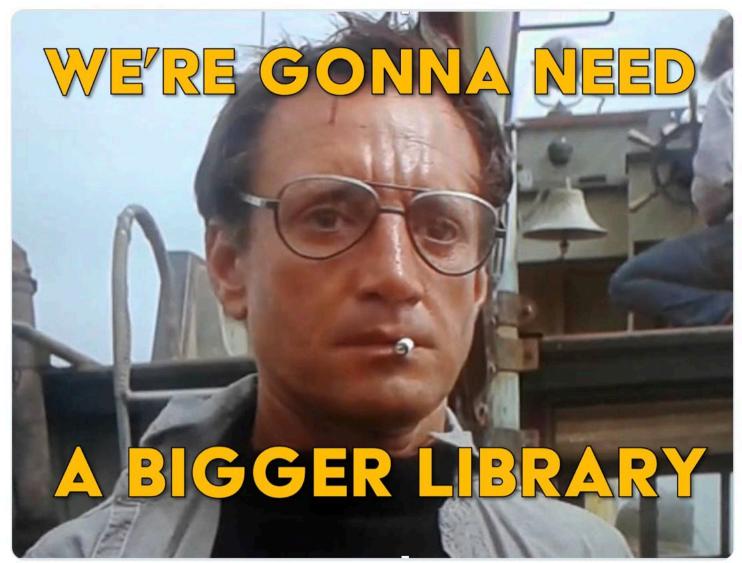
- examples of \overline{\
- examples of scholarly content that X scholarly record.
- **5** examples of content that's hard to archive, hard to cite and at risk of link rot.

Thanks to Web 2.0, this is just the tip of a growing iceberg of 'wild content'.





Toby Green wants your feedback on @Policy\_Commons @Toby...
One thing is clear . . .





Toby Green wants your feedback on @Policy\_Commons @Toby...

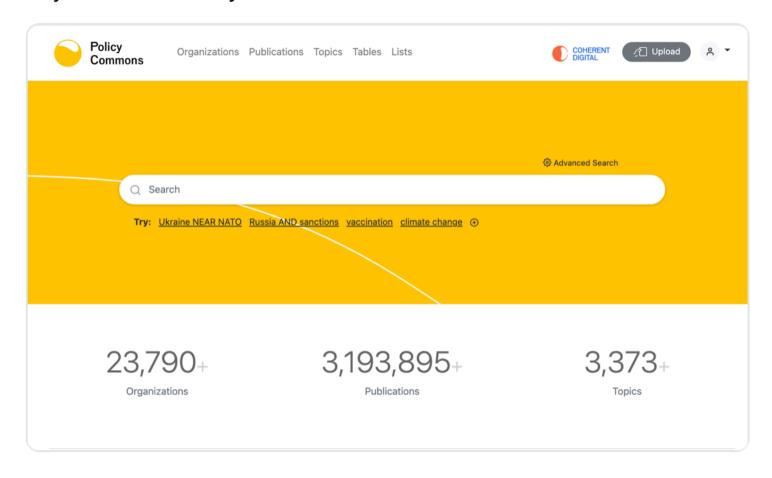
But how to capture all this 'wild content'?

'Combine' harvesting (ie harvesting, threshing etc in one process) by hand isn't going to work . . . we're going to need machines. Today, we're at the horse'n'harvest stage of 'combine information harvesting' . . .





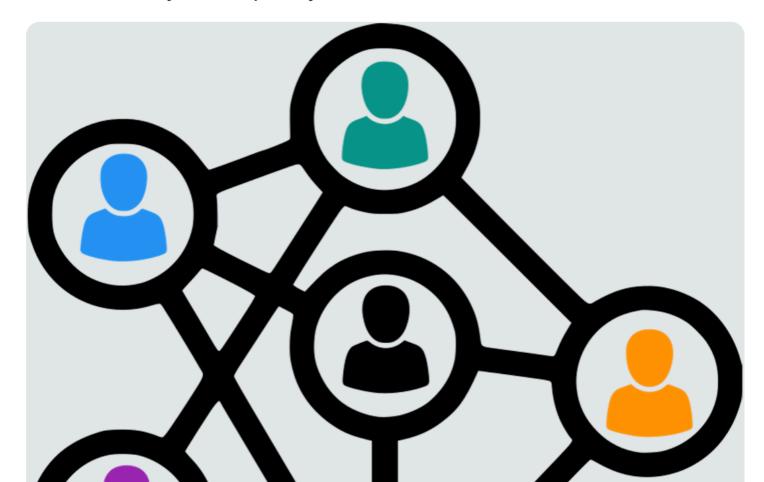
... but it's going well. @PolicyCommons has already harvested 3.2Mn items from 7,500 organisations (IGOs, NGOs, think tanks & research centres) - metadata'd and ready to get past the bouncer to join the scholarly record . . .





...including all [5] items mentioned in this [8]. You can check them out here https://policycommons.net/lists/271/fiesole-special-list/

But we need your help . . . join our network . . .





... you will know of organisations with 'wild content' that needs bringing into the scholarly record. We're ready when you are.

Thank you for reaching the end of this **I**!

Feedback and questions welcome.

to #Fiesole2022 for inviting me to create it.

